Revista Colombiana de Matemáticas
Volumen X (1976), págs. 57-68

AN ALGEBRAIC STUDY OF THE KOLMOGOROV ENTROPY

by

Steven M. MOORE

ABSTRACT

Some of the basic results on the Kolmogorov entropy are obtained within the formalism of C^*-álgebras, whether the C^*-algebra is commutative or not. Thus we have a theory which includes the classical and quantum systems.

RESUMEN

Algunos de los resultados básicos sobre la entropía de Kolmogorov son obtenidos dentro del formalismo de C^*-álgebras, siendo la C^*-álgebra conmutativa o no. Así obtenemos una teoría que incluye sistemas clásicos y cuánticos.

§ 1. ***Introduction.*** **In a previous article [1], we defined the C^*-algebra for classical and quantum systems. Consider an arbitrary C^*-algebra $\mathfrak{A}$ and a state ρ on $\mathfrak{A}$. The GNS construction gives us a cyclic representation π_ρ of $\mathfrak{A}$ in $\mathcal{B}(\mathcal{H}_\rho)$ with cyclic vector Ω_ρ such that**

$$\rho(A) = (\pi_\rho(A)\Omega_\rho, \Omega_\rho) \quad , \quad A \in \mathfrak{A} \quad . \tag{1}$$

The substitute for a probability measure is

$$r(P) = (P\,\Omega_\rho\ ,\ \Omega_\rho\ ) \tag{2}$$

where P is a projection in $\pi_\rho(\mathfrak{A})$. $\alpha = \{P_1, \ldots, P_n\}$ is a partition of Ω_ρ if the projections P_i are mutually orthogonal and $P_i\Omega_\rho \neq 0$ for $i = 1, \ldots, n$, with $\Sigma P_i\,\Omega_\rho = \Omega_\rho$. Note that we have slightly changed the definition of α used in [1]. Here we give emphasis to the fact that it is really the partitioning of Ω_ρ which is important. The extension of the Kolmogorov entropy of classical systems [2] for the partition α is defined in algebraic language by

$$h(\alpha) = -\Sigma\ r(P_i)\ \log_2 r(P_i)\ . \tag{3}$$

Note that we have abstracted away all reference to classical and quantum systems. In fact, we have just given a definition of the Kolmogorov entropy good for any C^*-algebra. If $\mathfrak{A}$ is a commutative C^*-algebra and the spectrum of $\mathfrak{A}$ is a compact C^∞- manifold, then $\mathfrak{A}$ corresponds to a classical system ; but for the main part of this paper, the exact structure of $\mathfrak{A}$ is not important.

We now introduce several simplifications in the notation. First, we eliminate the subscript ρ on π, $\mathcal{H}$ and Ω if no confusion is likely to arise. Next we introduce

$$I(t) = \begin{cases} -t\ \log_2 t & \text{if} \quad 0 < t \leq 1 \\ 0 & \text{if} \quad t = 0 \end{cases} \tag{4}$$

$I(t) \geq 0$, is continuous, and is strictly concave. $I(t) = 0$ if, and only if $t = 0$ or $t = 1$. Now (3) reduces to

$$h(\alpha) = \Sigma\ I(r(P_i)\ ) \tag{5}$$

In the next two sections we transcribe the classical treatment of the Kolmogorov entropy (see [3], for example) into the algebraic formalism.

§ 2. *Properties of* $h(\alpha)$.

2.1. *PROPOSITION.* a) $h(\alpha) = 0$ *if, and only if*, $\alpha = \{P\}$ *where* $P\Omega = \Omega$.

b) $h(\alpha) \leq \log_2 n$, *where* n *is the cardinality of* α *and* $h(\alpha) = \log_2 n$ *if, and only if,* $r(P_i) = 1/n$ *for each* i.

Proof. a) If $\alpha = \{P\}$, $P\Omega = \Omega$, then $r(P) = 1$, so $h(\alpha) = 0$. If $h(\alpha) = 0$, and $\alpha = \{P_1, \ldots, P_n\}$, then each $I(r(P_i)) = 0$. Thus $r(P_i)$ is 0 or 1, i.e. $(P_i\Omega, \Omega)$ is 0 or 1. Since $P_i\Omega \neq 0$, $(P_i\Omega, \Omega) = (P_i\Omega, P_i\Omega) \geq 0$, so there is only one projection P in α, and $P\Omega = \Omega$.

b) If $\alpha = \{P_1, \ldots, P_n\}$, by Jensen's inequality ([4] p. 61) we have

$$h(\alpha) = \Sigma\, I(r(P_i)) = n \Sigma \frac{1}{n} I(r(P_i))$$

$$\leq n\, I\left(\frac{1}{n} \Sigma\, r(P_i)\right) = n\, I\left(\frac{1}{n}\right) = \log_2 n$$

with equality if, and only if $r(P_i) = 1/n$. Q.E.D.

We say that $\alpha = \{P_i\}$, $\beta = \{Q_j\}$ are compatible if every projection P_i commutes with every projection Q_j. We write $\alpha \leq \beta$ if every subspace associated with β is contained in some subspace of α, i.e. $Q_j \leq P_{i_j}$ for some i_j for each j. Obviously $\alpha \leq \beta$ implies α and β are compatible. If α and β are compatible, we define

$$\alpha \vee \beta = \{P_i Q_j : P_i Q_j \Omega \neq 0\} \tag{6}$$

We obviously have $\alpha \leq \alpha \vee \beta$, $\beta \leq \alpha \vee \beta$. $\alpha \vee \beta$ is called the common refinement of α and β. It is easy to extend (6) to obtain $\alpha_1 \vee \ldots \vee \alpha_n$ where $\alpha_1, \ldots, \alpha_n$ are mutually compatible partitions. We note that compatibility is automatic when $\mathfrak{A}$ is commutative.

If $\alpha = \{P_i\}$ and $\beta = \{Q_j\}$ are compatible partitions, we define

$$h(\alpha \mid \beta) = \sum_{j=1}^{m} r(Q_j) \sum_{i=1}^{n} I(r(P_i \mid Q_j)) \tag{7}$$

where

$$r(P_i \mid Q_j) = \frac{r(P_i Q_j)}{r(Q_j)} \tag{8}$$

2.2. THEOREM. $b(\alpha \mid \beta)$ has the following properties :

$$b(\alpha \mid \beta) \geq 0 \tag{9}$$

with equality if, and only if , $\alpha \leq \beta$;

$$b(\alpha \vee \beta \mid \gamma) = b(\alpha \mid \gamma) + b(\beta \mid \alpha \vee \gamma) \ ; \tag{10}$$

$$\alpha \leq \beta \ \Rightarrow \ b(\alpha \mid \gamma) \leq b(\beta \mid \gamma) \ ; \tag{11}$$

$$\beta \leq \gamma \ \Rightarrow \ b(\alpha \mid \gamma) \geq b(\alpha \mid \beta) \ ; \tag{12}$$

$$b(\alpha \vee \beta \mid \gamma) \leq b(\alpha \mid \gamma) + b(\beta \mid \gamma) \ ; \tag{13}$$

Proof. (9) is obvious. Suppose $\alpha \leq \beta$, with $\alpha = \{P_i\}$ and $\beta = \{Q_j\}$. Since each Q_j in β satisfies $Q_j \leq P_{i_j}$ for some P_{i_j} in α, $P_{i_j} Q_j = Q_j$ while $P_i Q_j = 0$ for $i \neq i_j$. Thus

$$b(\alpha \mid \beta) = \sum_{j=1}^{m} r(Q_j) \sum_{i=1}^{n} I\left(\frac{r(P_i Q_j)}{r(Q_j)}\right) = \sum_{j=1}^{m} r(Q_j)\, I\left(\frac{r(Q_j)}{r(Q_j)}\right) = 0$$

Now suppose $b(\alpha \mid \beta) = 0$. Then

$$r(Q_j)\, I\left(\frac{r(P_i Q_j)}{r(Q_j)}\right) = 0$$

for each j and i. Since $r(Q_j) \neq 0$, we have that $r(P_i Q_j)$ is either equal to $r(Q_j)$ or equal to zero. Since $\sum_i r(P_i Q_j) = r(Q_j)$, there is only one index i_j such that $r(P_{i_j} Q_j) = r(Q_j)$ while $r(P_i Q_j) = 0$ for $i \neq i_j$. Thus $Q_j \leq P_{i_j}$.

Now let $\gamma = \{R_k\}$. The elements of $\alpha \vee \beta$ and $\alpha \vee \gamma$ are respectively $P_i Q_j$, $P_i R_k$. Thus

$$b(\alpha \vee \beta \mid \gamma) = -\sum_{ijk} r(P_i Q_j R_k) \log_2 r(P_i Q_j \mid R_k) .$$

But

$$r(P_iQ_j \mid R_k) = \frac{r(P_iR_k)}{r(R_k)} \quad , \quad \frac{r(P_iQ_jR_k)}{r(P_iR_k)} = r(P_i \mid R_k)\, r(Q_j \mid P_iR_k) \; ,$$

so that

$$h(\alpha \vee \beta \mid \gamma) = -\sum_{ijk} r(P_iQ_jR_k) \log_2 r(P_i \mid R_k) - \sum_{ijk} r(P_iQ_jR_k) \log_2 r(Q_j \mid P_i R_k)$$

$$= -\sum_{ik} r(P_iR_k) \log_2 r(P_i \mid R_k) - \sum_{ijk} r(Q_jP_iR_k) \log_2 r(Q_j \mid P_iR_k)$$

$$= h(\alpha \mid \gamma) + h(\beta \mid \alpha \vee \gamma) \; .$$

This gives us (10) .

If $\alpha \leq \beta$, then $\alpha \vee \beta = \beta$ and (9) and (10) imply

$$h(\beta \mid \gamma) = h(\alpha \mid \gamma) + h(\beta \mid \alpha \vee \gamma) \geq h(\alpha \mid \gamma) \; .$$

This gives us (11) .

Since $\sum_k r(R_k \mid Q_j) = 1$ and $r(R_k \mid Q_j) \geq 0$, the concavity of $I(t)$ implies that

$$\sum_k I(r(P_i \mid R_k))\, r(R_k \mid Q_j) \leq I\Big(\sum_k r(P_i \mid R_k)\; r(R_k \mid Q_j)\Big)$$

since $\beta \leq \gamma$, for each k, $Q_j \geq R_k$ for some j . Thus

$$\sum_k r(P_i \mid R_k) r(R_k \mid Q_j) = \sum_{k'} \frac{r(P_iR_{k'})\, r(R_{k'})}{r(R_{k'})\, r(Q_j)} = r(P_i \mid Q_j) \, ,$$

where the sum extends over those k' such that $R_{k'} \leq Q_j$ for some j . Thus we have

$$\sum_k I(r(P_i \mid R_k))\, r(R_k \mid Q_j) \leq I(r(P_i \mid Q_j)) \; .$$

Multiplying both sides by $r(Q_j)$ and summing over i and j yields equation (12).

Finally, (13) is a consequence of (10) and (12) : $\alpha \vee \gamma \geq \gamma$ implies that

$$h(\beta \mid \alpha \vee \gamma) \leq h(\beta \mid \gamma)$$

and

$$h(\alpha \vee \beta \mid \gamma) = h(\alpha \mid \gamma) + h(\beta \mid \alpha \vee \gamma) \leq h(\alpha \mid \gamma) + h(\beta \mid \gamma) .$$

Q. E. D.

2.3. *COROLLARY. $h(\alpha)$ has the following properties :*

$$h(\alpha \vee \beta) = h(\alpha) + h(\beta \mid \alpha) \; ; \qquad (14)$$

$$\alpha \leq \beta \Rightarrow h(\alpha) \leq h(\beta) \; ; \qquad (15)$$

$$h(\alpha \mid \beta) \leq h(\alpha) \; ; \qquad (16)$$

$$h(\alpha \vee \beta) \leq h(\alpha) + h(\beta) \; ; \qquad (17)$$

Proof. Take $\gamma = \{I\}$. Then $h(\alpha \mid \gamma) = h(\alpha)$, $\alpha \vee \gamma = \alpha$ since $\gamma \leq \alpha$, and the rest follows from the theorem . Q. E. D .

We note that it is probably possible to extend these definitions to denumerable partitions, but we will not need this extension here ([5] gives the extension in the classical case).

§ 3. *Entropy of a "measure preserving" transformation.* We now want to extend Kolmogorov's definition of entropy per unit time interval [6] to the algebraic case we are treating. Let $\{U_t\}$ be a one-parameter continuous group of unitary operators in $\mathfrak{A}$, where $U_t = exp(-iHt)$ and H is self-adjoint. We suppose that ρ is invariant with respect to $\{U_t\}$, i.e. $\rho(U_t^* A U_t) = \rho(A)$. Then there is a unitary representation $\{V_t\}$ of $\{U_t\}$ in $\mathcal{B}(\mathcal{H})$ so that for all t.

$$V_t^* \pi(A) V_t = \pi(U_t^* A U_t) \qquad (18)$$

$$V_t^* \Omega = \Omega \qquad (19)$$

If P is a projection, define $P_t = V_t^* P V_t$. Then $r(P_t) = r(P)$, so r is inva -

riant under V_t. Take $T = V_1$. Then r is also invariant under T. We call such a T a "measure preserving" transformation because it preserves the "measure" r.

3.1. LEMMA. a) *If* α *is a partition, so is* $T\alpha$.

b) $T(\alpha \vee \beta) = T\alpha \vee T\beta$

c) $h(\alpha \mid \beta) = h(T\alpha \mid T\beta)$

Proof. a) If P is a projection in $\pi(\mathfrak{A})$, then T^*PT is also a projection in $\pi(\mathfrak{A})$ since $T^*PT = \pi(U_1^* W U_1)$ if $P = \pi(W)$, by (18). Moreover $T\Omega = \Omega$, so $T^*PT\Omega = 0$ implies that $T^*P\Omega = 0$. Thus $0 = (T^*P\Omega, T^*P\Omega)$, or $0 = (P\Omega, P\Omega)$ which is a contradiction if P is an element of a partition.

b) We have

$$T(\alpha \vee \beta) = \{T^* P_i Q_j T : P_i Q_j \Omega \neq 0\} = \{T^* P_i T T^* Q_j T : T^* P_i T T^* Q_j T\Omega \neq 0\}$$

$$= T\alpha \vee T\beta$$

c) We have

$$h(\alpha \mid \beta) = \sum_{j=1}^{m} r(Q_j) \sum_{i=1}^{n} I\left(\frac{r(P_i Q_j)}{r(Q_j)}\right) = \sum_{j=1}^{m} r(T^* Q_j T) \sum_{i=1}^{n} I\left(\frac{r(T^* P_i T T^* Q_j T)}{r(T^* Q_j T)}\right)$$

$$= h(T\alpha \mid T\beta).$$

Q. E. D.

We now define the entropy of α relative to T by

$$h(\alpha, T) = \lim_{n \to \infty} \frac{h(\alpha \vee T\alpha \vee \ldots \vee T^{n-1}\alpha)}{n} \tag{20}$$

where we assume $T^k\alpha$ and $T^l\alpha$ are compatible for all integers k and l. We denote the class of all such α by $\Gamma(T)$.

3.2. THEOREM. The limit exists and is equal to

$$\lim_{n \to \infty} h(\alpha \mid T^{-1}\alpha \vee \ldots \vee T^{-n}\alpha)$$

Proof. We set

$$h_n = h(\alpha \vee T\alpha \vee \ldots \vee T^{n-1}\alpha),$$

$$s_n = h_n - h_{n-1}.$$

By (15) $\alpha \vee \ldots \vee T^{n-1}\alpha \leq \alpha \vee \ldots \vee T^n\alpha$ implies that $h_{n-1} \leq h_n$, so $s_n \geq 0$. But by (14) we have

$$s_n = h(\alpha \vee \ldots \vee T^{n-1}\alpha) - h(\alpha \vee \ldots \vee T^{n-2}\alpha)$$
$$= h(T^{n-1}\alpha \mid \alpha \vee \ldots \vee T^{n-2}\alpha).$$

Hence

$$s_{n-1} = h(T^{n-2}\alpha \mid \alpha \vee \ldots \vee T^{n-3}\alpha),$$

so by the lemma

$$s_{n-1} = h(T^{n-1}\alpha \mid T\alpha \vee \ldots \vee T^{n-2}\alpha).$$

Since $T\alpha \vee \ldots \vee T^{n-2}\alpha \leq \alpha \vee \ldots \vee T^{n-2}\alpha$, (12) gives us $s_{n-1} \geq s_n$. Thus $\{s_n\}$ is a decreasing sequence of positive numbers, so $s_n \to s$ for some $s \geq 0$. Since

$$h_n = h(\alpha) + s_1 + \ldots + s_n,$$

Cesaro's mean convergence theorem implies that

$$\lim_{n\to\infty} \frac{h_n}{n} = s.$$

Q.E.D.

We mow define (as an extension of the definition in [6]) the entropy of the "measure preserving" transformation T by

$$h(T) = \sup_{\alpha \in \Gamma(T)} h(\alpha, T) \qquad (21)$$

i.e. $h(T)$ is the sup over all α such that $T^k\alpha$ and $T^\ell\alpha$ are compatible for all integers k and ℓ.

§ 4. *Spectral properties of* $h(T)$. We say that the triple $(\mathfrak{A}, \rho, T)$ is spectrally equivalent to the triple $(\mathfrak{A}', \rho', T')$ if there is a $*$-isomorphism $\varphi : \mathfrak{A} \to \mathfrak{A}'$ such that :

$$\varphi (U_t^* A U_t) = U_t'^{*} \varphi (A) U_t' \tag{22}$$

$$(\pi_\rho (A) \Omega_\rho, \Omega_\rho) = (\pi_{\rho'} (\varphi (A)) \Omega_{\rho'}, \Omega_{\rho'}) \tag{23}$$

In this case we write $(\mathfrak{A}, \rho, T) \simeq (\mathfrak{A}', \rho', T')$. We note that (22) says that the time development of $\mathfrak{A}$ is equivalent to that of $\mathfrak{A}'$ if we interpret the parameter t as time. (23) says that the expectation values are conserved.

4.1 THEOREM. *Let* $(\mathfrak{A}, \rho, T) \simeq (\mathfrak{A}', \rho', T')$. *Then there exists a unitary operator* $V : \mathcal{H}_\rho \to \mathcal{H}_{\rho'}$ *such that* :

$$V \pi_\rho (A) V^* = \pi_{\rho'} (\varphi (A)) \tag{24}$$

$$V \Omega_\rho = \Omega_{\rho'} \tag{25}$$

$$V T V^* = T' \tag{26}$$

Proof. Define V on the dense set $\{ \pi_\rho (A) \Omega_\rho \}$ by

$$V \pi_\rho (A) \Omega_\rho = \pi_{\rho'} (\varphi (A)) \Omega_{\rho'}, \quad \text{for all} \quad A \in \mathfrak{A} .$$

We have

$$(V \pi_\rho (A) \Omega_\rho, V \pi_\rho (B) \Omega_\rho) = (\pi_{\rho'} (\varphi (A)) \Omega_{\rho'}, \pi_{\rho'} (\varphi (B)) \Omega_{\rho'})$$

$$= (\pi_{\rho'} (\varphi (B^* A)) \Omega_{\rho'}, \Omega_{\rho'}) = (\pi_\rho (B^* A) \Omega_\rho, \Omega_\rho)$$

$$= (\pi_\rho (A) \Omega_\rho, \pi_\rho (B) \Omega_\rho),$$

so V is an isometry. (24) and (25) are obvious from the definition of V. For (26), we recall that the canonical construction of T is

$$T \pi_\rho (A) \Omega_\rho = \pi_\rho (U_1^* A U_1) \Omega_\rho .$$

Similarly for T'. Thus

$$V T V^* \pi_{\rho'}(A') \Omega_{\rho'} = V T \pi_\rho(\varphi^{-1}(A')) \Omega_\rho$$

$$= V \pi_\rho(U_1^* \varphi^{-1}(A') U_1) \Omega_\rho$$

$$= \pi_{\rho'}(\varphi(U_1^* \varphi^{-1}(A') U_1)) \Omega_{\rho'}$$

$$= \pi_{\rho'}(U_1'^* A' U_1') \Omega_{\rho'}$$

$$= T' \pi_{\rho'}(A') \Omega_{\rho'} .$$

Q.E.D.

4.2. THEOREM. The Kolmogorov entropy is an invariant for spectrally equivalent systems.

Proof. Let $(\mathfrak{A}, \rho, T) \simeq (\mathfrak{A}', \rho', T')$, α, and V be as above. Then, if α is a partition of Ω_ρ, $V\alpha$ is a partition of $\Omega_{\rho'}$, so

$$h(V\alpha, T') = h(V\alpha, VTV^*) = \lim \frac{h(V\alpha \vee \ldots \vee (VTV^*)^n V\alpha)}{n}$$

$$= \lim \frac{h(V\alpha \vee VTV^* V\alpha \vee \ldots \vee VT^n V^* \alpha)}{n}$$

$$= \lim \frac{h(V(\alpha \vee T\alpha \vee \ldots \vee T^n \alpha))}{n}$$

$$= h(\alpha, T)$$

Thus $h(T') = h(T)$. Q. E. D.

§ 5. *The Kolmogorov entropy for two special C^*-algebras.* One of the nontrivial results about the Kolmogorov entropy of a classical system is the following theorem of Kuchnirenko [7] :

5.1. THEOREM. For a bounded classical system $h(T)$ is finite.

The generalization of this theorem to an arbitrary C^*- algebra does not seem

to be an easy task. Even in the commutative case, we need much more than mere commutativity. If $\mathfrak{A}$ is a commutative C^*-algebra with identity, then its spectrum is automatically compact T_2, but in general it can be very "unsmooth". (See [8] p. 273 for an example of extreme pathological behavior). The proof of Kuchnirenko's theorem uses very strongly the fact that the spectrum of a bounded classical system is a compact C^∞- manifold. In any case, we wish to give a rather trivial example of this theorem in a quantum mechanics context which covers finite lattices or a system of a finite number of spins :

5.2. THEOREM. If $\mathcal{H}_\rho$ is finite-dimensional, $h(T)$ is finite.

Proof. Let dim $\mathcal{H}_\rho = N$. If $N = 1$, $h(\alpha) = 0$ for any α, so the case $N = 1$ is trivial. Assume $N > 1$. From Proposition 2.1 we have that $h(\alpha) \leq \log_2 n$ where n is the cardinality of α. But $n \leq N$, so $h(\alpha) \leq \log_2 N \leq N$. By Corollary 2.3 and Lemma 3.1

$$h(\alpha \vee T\alpha \vee \ldots \vee T^{k-1}\alpha) \leq \sum_{i=o}^{k-1} h(T^i\alpha) = \sum_{i=o}^{k-1} h(\alpha) \leq Nk$$

Thus $h(\alpha, T) \leq N$. Q.E.D.

We wish to propose the following question as an unsolved problem to this date: What are the weakest conditions on $\mathfrak{A}$ that guarantee that $h(T)$ is finite ?

References

1. *S. M. Moore : An Algebraic Formulation of Ergodic Problems.* Rev. Col. Mat. 9, 189-204 (1975).
2. *A. N. Kolmogorov : A New Metric Invariant of Transitive Dynamic Systems and Automorphisms of Lebesgue Fields.* Doklady Akad. Nauk 119 No. 5, 861-864 (1958).
3. *V. I. Arnold, A. Avez : Ergodic Problems of Classical Mechanics.* W.A. Benjamin, New York, 1968.
4. *W. Rudin : Real and Complex Analysis.* McGraw-Hill, New York, 1966.
5. *V.A. Roblin, Ya. Sinai : Construction and Properties of Invariant Measurable*

Partitions. Doklady Akad. Nauk. 141 No. 6, 1038-1041 (1961). (Also in English : Sov. Math. Doklady 2 No. 6, 1611-1614 (1961)).

6. *A. N. Kolmogorov : On the Entropy per Unit Time as a Metric Invariant of Automorphisms.* Doklady Akad. Nauk. 124 No. 4, 754-755 (1959).

7. *A. G. Kuchnirenko : An Estimate from Above for the Entropy of a Classical System.* Doklady Akad. Nauk. 161 No. 1, 37-38 (1965). (Also in English :Sov. Math. Doklady 6 No. 2, 360-362 (1965).)

8. *W. Rudin : Functional Analysis,* McGraw-Hill, New York, 1973.

Departamento de Física
Universidad Pedagógica Nacional
Bogotá, D.E., Colombia, S. A.

(Recibido en agosto de 1975).